



**Barcelona  
Supercomputing  
Center**

*Centro Nacional de Supercomputación*



EXCELENCIA  
SEVERO  
OCHOA

# Gestión de datos en Ciencias de la Tierra

Francesco Benincasa <francesco.benincasa@bsc.es>



## **El Departamento**

- composición

## **Datos en ciencias de la tierra**

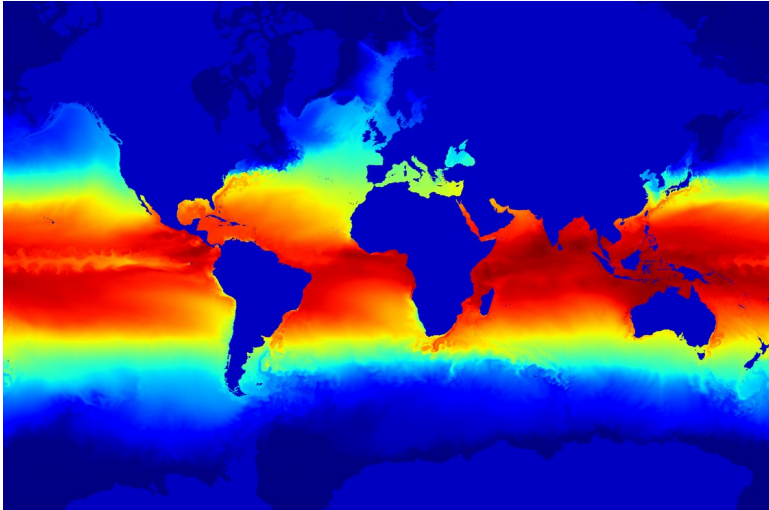
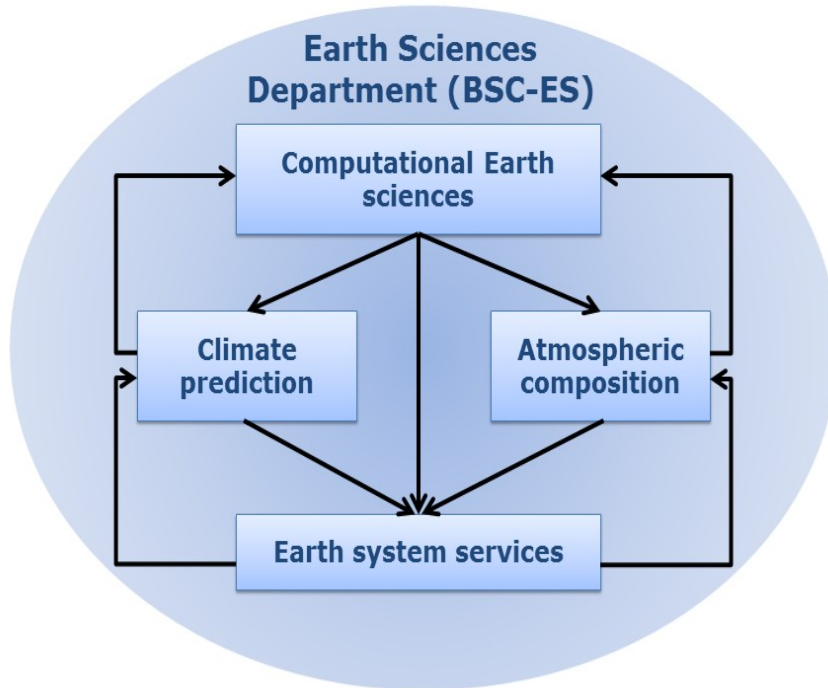
- tipos de datos
- fuentes de datos
- formatos: NetCDF/HDF/GRIB
- organización interna de los datos

## **Sistemas de gestión de datos**

- proyectos puntuales
- proyectos a grande escala



# El Departamento



## Atmospheric composition

- predicciones de tiempo, calidad del aire y polvo mineral a corto plazo

## Climate prediction

- predicciones climáticas a medio-largo plazo, desde estaciones hasta décadas

## Computational earth sciences

- mejoramiento de código de modelos, gestión de datos y desarrollo de software

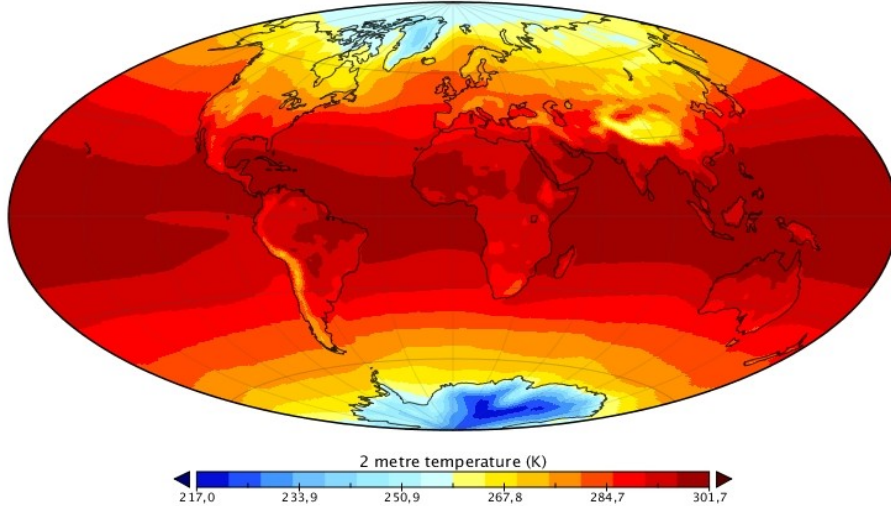
## Earth system services

- desarrollo de servicios y estudios de clima y calidad del aire en colaboración con administraciones públicas y para empresas



# Datos en ciencias de la tierra

Mean Surface Temperature (1990–1999) for EC-EARTH at ES-BSC



## Variables 4-dimensionales

- latitud
- longitud
- pasos de tiempo
- altitud/nivel vertical o experimento

## Variables vectoriales

- viento

## Variables calculadas

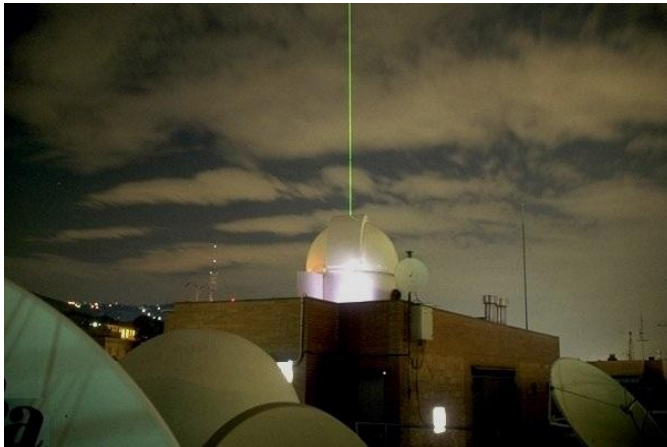
- acumulaciones

## Sistemas de coordenadas

- regular
- curvilínea (varios tipos)







## Modelos numéricos

- desarrollados internamente (NMMB/BSC y BSC-Dream8b)
- externos a los que colaboramos (EC-Earth o NEMO)
- externos

## Observaciones

- redes de estaciones internacionales
- redes de estaciones nacionales y autonómicas
- estaciones locales
- boyas
- satélites

[Record 0]

Center=7

Process=81

GridID=255

Parameter=U GRD

Layer=105/10

Date=2011/12/04 00:00

Valid Time=0

LatPts=11

LonPts=21

Lat=45°00'N,50°00'N

dLat=0.50

Lon=350°00'E,000°00'E

dLon=0.50

Flags=11

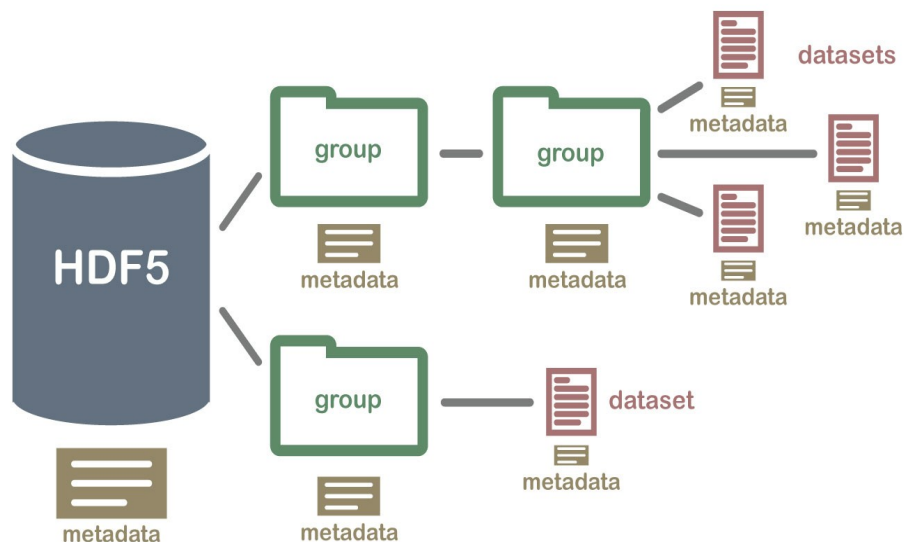
Data Values:

10.2, 10.3, 10.5, 10.7, 10.9, 11.2, 11.5, 11.8, 11.9, 12.1, 12.1, 11.9, 11.8, 11.6, 11.3, 10.9,  
11.4, 11.5, 11.6, 11.7, 11.8, 11.8, 11.6, 11.6, 11.6, 11.6, 11.5, 11.3, 11.1, 10.9, 10.6, 10.3,  
12.0, 11.9, 11.9, 11.9, 12.1, 12.1, 11.8, 11.6, 11.4, 11.3, 11.0, 10.7, 10.4, 10.2, 9.9, 9.8,  
11.7, 11.7, 11.8, 11.8, 11.9, 11.8, 11.6, 11.5, 11.4, 11.1, 10.7, 10.4, 10.2, 9.9, 10.0, 10.2,  
11.3, 11.2, 11.3, 11.3, 11.3, 11.0, 10.7, 10.5, 10.4, 10.3, 10.0, 9.8, 9.8, 9.9, 10.4, 9.6,  
10.8, 10.9, 10.9, 10.8, 10.6, 10.4, 10.3, 10.1, 9.8, 9.5, 9.3, 9.3, 9.0, 8.3, 8.2, 6.9, 4.7,  
10.7, 10.8, 10.9, 10.8, 10.7, 10.6, 10.4, 10.2, 10.0, 9.7, 9.2, 6.8, 4.1, 2.9, 3.1, 3.7, 3.9,  
10.8, 11.0, 11.1, 11.2, 11.2, 11.1, 10.8, 10.6, 10.4, 10.2, 9.7, 7.4, 4.9, 4.0, 4.7, 5.9,  
10.3, 10.4, 10.6, 10.9, 11.1, 11.1, 10.9, 10.6, 10.3, 10.0, 9.9, 9.8, 9.9, 10.1, 10.8, 12.1,  
9.1, 9.4, 9.9, 10.4, 10.8, 10.9, 10.7, 10.4, 9.9, 9.6, 9.5, 9.8, 10.5, 11.5, 12.2, 12.8,  
9.1, 9.4, 9.9, 10.4, 10.8, 10.7, 10.2, 9.6, 9.2, 8.7, 8.5, 9.0, 10.0, 11.1, 12.0, 12.7, 13.1,

## GRidded Information in Binary

- aprobado por la **OMM/WMO** (Organización Meteorológica Mundial)
- utilizado por los centros nacionales de predicción (AEMET) y los centros internacionales como EUMETSAT, ECMWF o NCEP
- codificado por tablas publicas (cada centro y cada variable identificada por un numero)
- conjunto de registros **auto-contenidos** de datos 2D
- utilizado por nosotros para intercambiar datos con instituciones nacionales e internacionales





```
HDF5 "test.h5" {  
  GROUP "/" {  
    GROUP "Group1" {  
      DATASET "Dataset1" {  
        DATATYPE { H5T_STD_U32LE }  
        DATASPACE { SIMPLE ( 4 ) / ( 4 ) }  
        DATA { 0, 3, 6, 9 }  
      }  
      DATASET "Dataset2" {  
        DATATYPE { H5T_STD_U8LE }  
        DATASPACE { SIMPLE ( 4 ) / ( 4 ) }  
        DATA { 0, 0, 0, 0 }  
      }  
    }  
  }  
}
```

## Hierarchical Data Format

- creado por la NCSA, ahora desarrollado y mantenido por el HDF Group
- parecido a un file-system
- **group**: carpeta
- **dataset**: el dato
- cada elemento tiene meta-datos asociados
- puede contener datos heterogéneos
- incluye compresión
- proporciona la posibilidad de extracción parcial de datos

dimensions:

```
time = 25 ;  
lat = 347 ;  
lon = 423 ;
```

variables:

```
float SCONC_DUST(time, lat, lon) ;  
    SCONC_DUST:_FillValue = 9.96921e+36f ;  
    SCONC_DUST:long_name = "Surface Dust Concentration" ;  
    SCONC_DUST:units = "kg/m3" ;
```

```
float OD550_DUST(time, lat, lon) ;  
    OD550_DUST:_FillValue = 9.96921e+36f ;  
    OD550_DUST:long_name = "AOD" ;  
    OD550_DUST:units = "-" ;
```

```
float time(time) ;  
    time:units = "hours since 2016-01-01 12:00:00" ;
```

```
float lon(lon) ;  
    lon:units = "degrees_east" ;
```

```
float lat(lat) ;  
    lat:units = "degrees_north" ;
```

// global attributes:

```
:history = "Tue Jan 19 16:47:47 2016: ncatted -a  
units,SCONC_DUST,o,c,kg/m3 2016010112_NOA-WRF-CHEM.nc\n",  
    "Tue Jan 19 16:47:47 2016: ncap2 -v -s  
SCONC_DUST=SCONC_DUST/1000000000; OD550_DUST=OD550_DUST  
2016010112_NOA-WRF-CHEM.nc_BAK 2016010112_NOA-WRF-  
CHEM.nc" ;  
:nco_openmp_thread_number = 1 ;
```

## Network Common Data Form

- desarrollado y mantenido por la Unidata (UCAR)
- incluye meta-datos
- conceptos de *dimensiones y variables*
- formato de referencia dentro del departamento

```
netcdf vas_20091101 {  
  dimensions:  
    ensemble = 5 ;  
    string4 = 4 ;  
    string15 = 15 ;  
    latitude = 256 ;  
    time = UNLIMITED ; // (4 currently)  
    longitude = 512 ;  
    string60 = 60 ;  
    time_bnd = 2 ;  
  
  variables:  
    char experiment_id(ensemble, string4) ;  
      experiment_id:standard_name = "experiment_id" ;  
      experiment_id:long_name = "Experiment identifier" ;  
    double latitude(latitude) ;  
      latitude:long_name = "latitude" ;  
      latitude:units = "degrees_north" ;  
      latitude:standard_name = "latitude" ;  
      latitude:axis = "Y" ;  
    float vas(time, ensemble, latitude, longitude) ;  
      vas:long_name = "near-surface meridional wind positive to the north"  
      ;  
      vas:grid_type = "gaussian" ;  
      vas:_FillValue = 1.e+12f ;  
      vas:standard_name = "northward_wind" ;  
      vas:cell_methods = "time: mean (interval 3 hours)" ;  
      vas:unit_long = "meter_per_second" ;  
      vas:units = "m s-1" ;  
      vas:coordinates = "longitude latitude reftime leadtime time_bnd  
      experiment_id source realization institution level_vas" ;  
}
```

## File system compartido

- separación por tipo de fuente (experimento, observaciones o reanalysis)
- separación por modelo
- separación por frecuencia
- separación por variable
- formato NetCDF

*//esnas/exp/ecearth/m00n/monthly\_  
mean/vas\_6hourly/vas\_200911.nc*



# Sistemas de gestión de datos

# CALIOPE sistema de predicción de calidad del aire

## CALIOPE

### Meteorology

- WRF-ARWv3.5
- 38 sigma levels (top 50 hPa)
- IBC: GFS (NCEP)
- 33 layers/50 hPa

### Emissions

- HERMESv2
- EU: HERMES-DIS (EMEP data)
- Spain: HERMES-BOUP

### Chemistry

- CMAQv5.0.1
- CB05/AERO5
- BC: NCAR MOZART4
- 15 layers/ 50 hPa

### Mineral Dust

- BSC-DREAM8bv2
- PM10 and PM2.5 desértico

### Post-processing

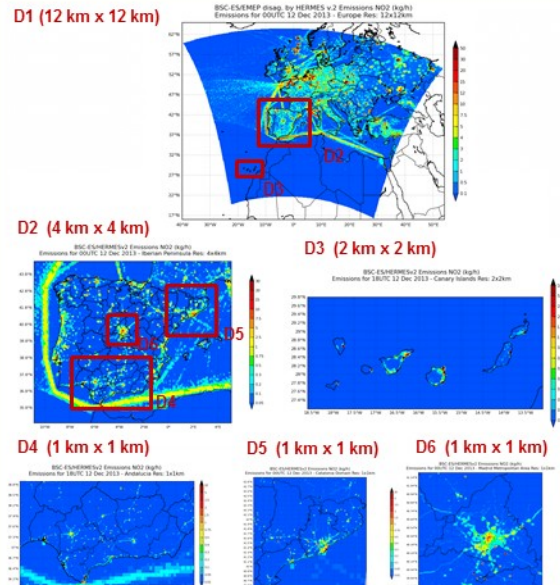
- Filtro de Kalman (puntual y 2D)

## Air quality predictions

O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, CO, PM10, PM2.5, Benceno

## 48h Forecast

- Concentration maps, emis, meteo, etc.
- Air quality index



## Evaluation (NRT)

- IONET
- Satellites



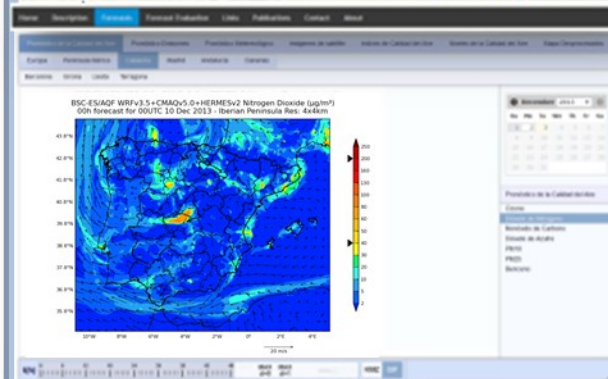
## Diffusion

- Web ([www.bsc.es/caliope](http://www.bsc.es/caliope))
- MyGEOSS



## Sistema CALIOPE

Pronóstico de la Calidad del Aire







## NEWSLETTER

Keep up to date with our activities!

Full Name

Your email

## SEARCH

Search Site

## HOME

- > About us
- > Forecast
- > Evaluation
- > Methods
- > News
- > Events
- > Contact

## LATEST NEWS

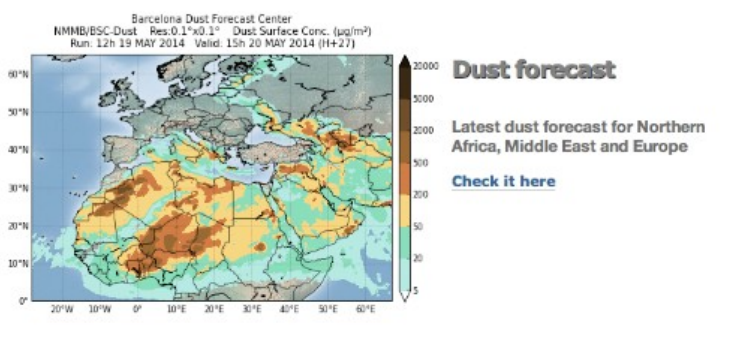
Service down for maintenance  
27-28 April  
Apr 23, 2014

Activity Report 2010-2012 of  
the SDS-WAS Regional Center  
published by the WMO  
Mar 25, 2014

## Barcelona Dust Forecast Center starts operations

The Center will release operational dust forecasts for Northern Africa, Middle East and Europe

[Read More](#)



- *Primer centro reconocido por la WMO para la predicción de polvo mineral*
- *Predicciones basadas en el modelo NMMB/BSC-Dust a 0.1° de resolución*
- *Proporciona imágenes y datos a WMO GTS, EumetCAST, AEMET, ...*
- *Evaluación en NRT*

**Colaboración entre el  
BSC y AEMET**







- Home**
- About us
  - Forecast & Products
  - Projects & Research
  - Materials
  - News
  - Events
  - Public Newsletter
  - Users Newsletter
  - SDS-WAS Survey

**Search**

Search Site Search

## Latest News

Released article introducing a new climatological index that modulates dust activity over SW Asia  
Jan 13, 2016

Materials from the workshop on the health impact of dust held in Amman  
Dec 17, 2015

National Observatory of Athens contributes to the SDS-WAS model intercomparison  
Dec 17, 2015

## Upcoming Events

34th National & 2nd International Geosciences

You are here: Home

## Northern Africa-Middle East-Europe (NA-ME-E) Regional Center

by Francesco Benincasa — last modified May 29, 2012 03:33 PM

### Outstanding

Latency in the provision of numerical forecasts

WMO Bulletin: Airborne Dust article

Kick-off meeting of the WMO SDS-WAS Steering Committee

The edge of crisis: Dust and sand storms

New product to monitor dust events

### Subscribe to the Public Newsletter!

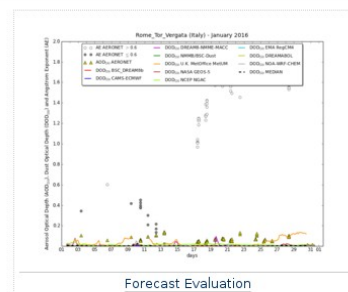
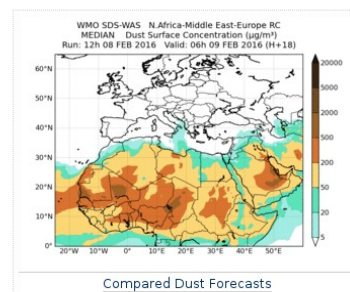
To be informed about our activities, news and events related to dust. Frequency is almost monthly.

Subscribe

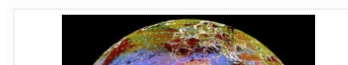
### Portal manual

Please find a brief manual here.

### Dust forecasts



### Dust observations



- Proyecto de investigación
- Conjunto de modelos proporcionado por partners internacionales
- Inter-comparación entre modelos
- Productos multi-modelo
- Evaluación AERONET
- Evaluación satélites
- Cálculo de estadísticas (BIAS, ...)
- Descarga de datos NetCDF

**Colaboración entre el  
BSC y AEMET**

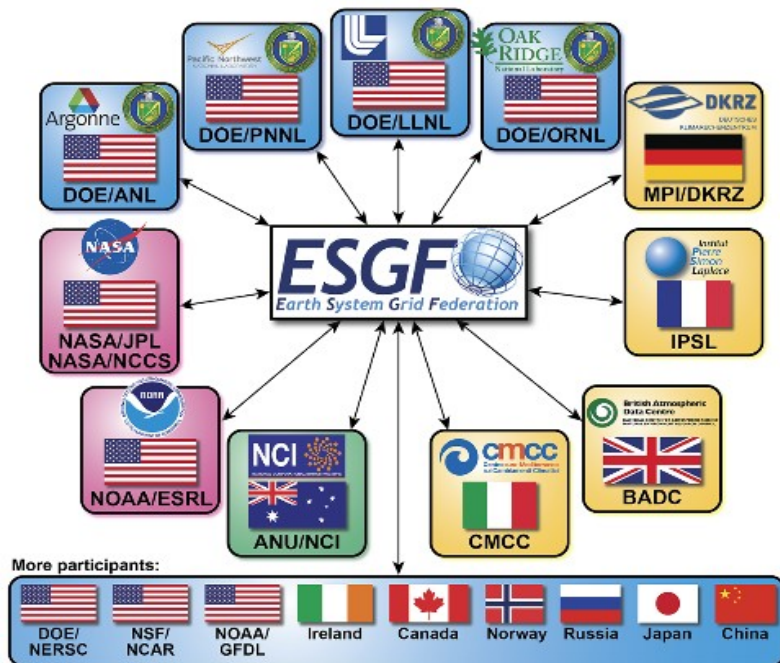
# Earth System Grid Federation



Barcelona  
Supercomputing  
Center  
Centro Nacional de Supercomputación

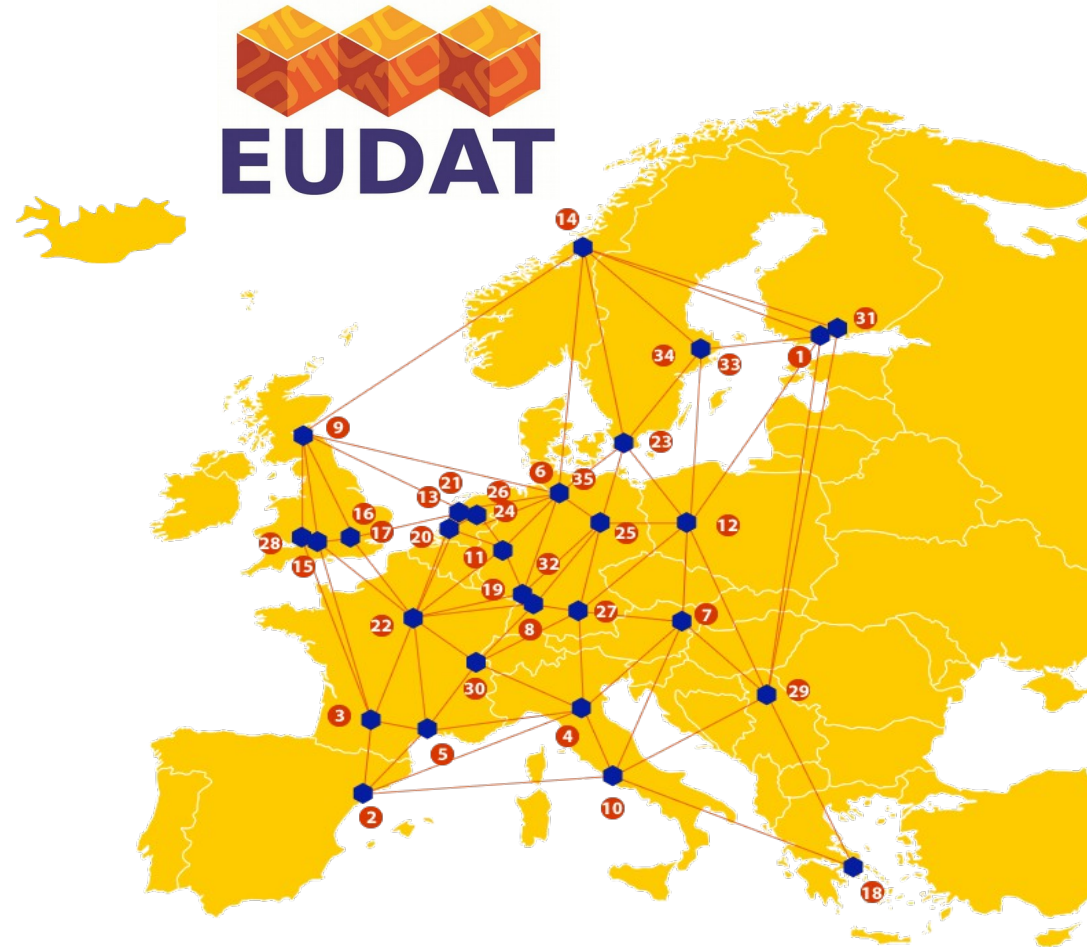
EXCELENCIA  
SEVERO  
OCHOA

*Plataforma “open source” robusta, distribuida para datos y calculo, que proporciona el acceso a una gran cantidad de datos científicos en todo el mundo.*



- Organizado como conjunto de nodos distribuidos en todo el mundo
- Data nodes (THREDDS server, gridFTP server, ...)
- Index nodes (búsqueda)
- Identity nodes (identificación)
- Compute nodes (análisis y visualización de datos)

- *Red Europea multidisciplinaria*
- *Servicios:*
  - Data sync & exchange
  - Data replication
  - Data discovery & search
  - Data repository & sharing
  - Data staging
  - ... more on work in progress





**Barcelona  
Supercomputing  
Center**

*Centro Nacional de Supercomputación*



¡Gracias!

[francesco.benincasa@bsc.es](mailto:francesco.benincasa@bsc.es)